Human Immunology 72 (2011) 426–430

Contents lists available at ScienceDirect

journal homepage: www.elsevier.com/locate/humimm

Genetic susceptibility to tuberculosis associated with cathepsin Z haplotype in a Ugandan household contact study

Allison R. Baker [a,†], Sarah Zalwango [b], LaShaunda L. Malone [c], Robert P. Igo Jr [a], Feiyou Qiu [a], Mary Nsereko [b], Mark D. Adams [c,d], Pamela Supelak [d], Harriet Mayanja-Kizza [b,e], W. Henry Boom [b,c], Catherine M. Stein [a,b,c,*]

[a] Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, Ohio, USA
[b] Uganda–CWRU Research Collaboration, Kampala, Uganda
[c] Tuberculosis Research Unit, Department of Medicine, Case Western Reserve University, Cleveland, Ohio, USA
[d] Department of Genetics and Center for Proteomics and Bioinformatics, Case Western Reserve University, Cleveland, Ohio, USA
[e] Makerere University School of Medicine and Mulago Hospital, Kampala, Uganda

ARTICLE INFO

Article history:
Received 20 November 2010
Accepted 22 February 2011
Available online 25 February 2011

Keywords:
Infectious disease
Family study
TB genetics
Fine map
Immunogenetics

ABSTRACT

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (*Mtb*), causes 9 million new cases worldwide and 2 million deaths annually. Genetic linkage and association analyses have suggested several chromosomal regions and candidate genes involved in TB susceptibility. This study examines the association of TB disease susceptibility with a selection of biologically relevant genes on regions on chromosomes 7 (*IL6* and *CARD11*) and 20 (*CTSZ* and *MC3R*) and fine mapping of the chromosome 7p22–p21 region identified through our genome scan. We analyzed 565 individuals from Kampala, Uganda, who were previously included in our genome-wide linkage scan. Association analyses were conducted for 1,417 single-nucleotide polymorphisms (SNP) that passed quality control. None of the candidate gene or fine mapping SNPs was significantly associated with TB susceptibility ($p > 0.10$). When we restricted the analysis to HIV-negative individuals, 2 SNPs on chromosome 7 were significantly associated with TB susceptibility ($p < 0.05$). Haplotype analyses identified a significant risk haplotype in cathepsin X (*CTSZ*; $p = 0.0281$, odds ratio = 1.5493, 95% confidence interval [1.039, 2.320]).

1. Introduction

Tuberculosis (TB) is an infectious disease caused by the bacterium *Mycobacterium tuberculosis* (*Mtb*). The World Health Organization reports that over 9 million new cases of TB are estimated each year, resulting in approximately 2 million deaths worldwide [1]. Only 5 to 10% of individuals infected with *Mtb* actually advance to disease [2]. The pathogenesis of TB follows a 2-stage process: a productive infection of *Mtb* whereby symptoms do not develop, followed by *Mtb* replication and the expression of disease symptoms [3]. TB disease is characterized by growth of *Mtb* on culture, presence of cavities on chest x-ray, and symptoms such as cough and fever.

Several arguments have been made for a genetic risk factor in development of TB, based on the observation that only a small percentage of individuals infected with *Mtb* progress to disease development. Although results remain somewhat inconsistent, animal models [4,5], twin studies [6–8], segregation analysis [9], candidate gene studies (reviewed in [10]), and linkage analysis [11–15] have all demonstrated evidence in support of a genetic component in the risk for developing TB. Stein et al. [16] conducted a genome-wide linkage analysis of Ugandan individuals, including both HIV-negative and HIV-positive individuals. Suggestive linkage to TB disease was found on a 34-cM-long segment on chromosome 7 ($p = 0.0002$), in addition to a 25-cM-long region on chromosome 20 ($p = 0.002$) [16]. Chromosome 7p22–7p21 contains the *IL6* and *CARD11* genes. Interleukin 6 (IL-6) is an immunoregulatory cytokine that inhibits production of tumor necrosis factor α (TNF-α) and IL-1β and thus may have a role in the response to mycobacterial infections [17]. IL-6-deficient mice succumb to TB infection, whereas wild-type mice do not [18]. However, Oral et al. [19] did not report significant differences in the distribution of the *IL6* gene polymorphisms or differences in IL-6 allele frequencies between human TB cases and controls. Such contradictory evidence merits further pursuit of the *IL6* gene's involvement in TB susceptibility. Further upstream in this chromosome 7 region is the gene *CARD11* (caspase recruitment domain family, member 11), which is part of the NOD-like receptor pathway. This gene is of interest because NOD-like receptors have nonredundant roles in *Mtb* recognition [20]. Although both *IL6* and *CARD11* are promising candidates, it is

* Corresponding author.
E-mail address: catherine.stein@case.edu (C.M. Stein).
† Current address: Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, Tennessee, USA.

0198-8859/11/$32.00 - see front matter
doi:10.1016/j.humimm.2011.02.016

possible that a novel locus exists under the linkage peak reported by Stein et al. [16].

The same region on chromosome 20q13 observed by Stein et al. [16] was reported to be a major susceptibility locus for TB in a study of South African and Malawian sibling pairs, HIV-negative and HIV-positive cases included [21]. Two genes in this chromosome 20 region, melanocortin 3 receptor (*MC3R*) and cathepsin Z (*CTSZ*), were mapped in South African and Malawian populations; these presented new candidate genes for TB [21]. Both chromosome 20 genes are biologically relevant to TB susceptibility because *MC3R* plays a suggested role in the regulation of energy homeostasis, whereas *CTSZ* is expressed in cancer cell lines with possible involvement in host defense and tumorigenesis [22].

The twin goals of this study were (1) to examine the association of TB susceptibility and the biologically relevant genes *IL6, CARD11, MC3R,* and *CTSZ*; and (2) to fine map the 1-LOD significance region on chromosome 7 identified through our previous genome scan.

2. Subjects and methods

2.1. Data description

Study participants were enrolled in the Household Contact Study (Phase I [23]) between 1995 and 1999 and in the Kawempe Community Health Study between 2002 and 2004 (Phase II [24]) in Kampala, Uganda. The individuals included in this analysis were analyzed in our genome scan [16] and had sufficient DNA remaining for additional genotyping. Of the individuals reported in that study, 18 have since developed culture-confirmed TB and 1 has since developed HIV infection; their trait data were recoded accordingly for this analysis.

A complete summary of study design and enrollment procedures can be found in our previous work [16,23,24]. Briefly, households were identified through an index case (proband) with culture-confirmed TB. All individuals within the household who provided informed consent were included in the study; individuals were not excluded based on HIV infection or relationship to the index case. All participants received a full clinical examination, including HIV testing, tuberculin skin test (TST), and sputum collection for TB suspects. Individuals with confirmed TB disease were treated appropriately [25]; in addition, individuals with latent *Mtb* infection were also offered preventive therapy [26]. Households were followed actively for 2 years after enrollment; after that 2-year period, individuals who developed symptoms of TB returned to the clinic for further evaluation and treatment. One of the strengths of this prospective design is the ability to observe all clinically relevant events, including both primary progressive and reactivation disease, as well as TST conversion. This long duration of follow-up significantly reduces the chances of misclassifying cases as controls [27]. This family-based design also has several advantages for genetic association studies [28]; in addition to addressing genetic epidemiologic concerns, such as population stratification, the family-based design enables the characterization of exposure to an infectious TB case [29].

2.2. Genotyping

Genotyping was performed using a custom Illumina GoldenGate single nucleotide polymorphism (SNP) panel comprising 1,536 SNPs with 1,528 from chromosome 7 (1,479 in the fine mapping region) and 8 from chromosome 20 (Supplemental Table 1). We first selected SNPs in our 4 candidate genes of interest: *IL6, CARD11, CTSZ*, and *MC3R*. Haplotype tagging SNPs were selected for *IL6, CTSZ*, and *MC3R* using all 3 African HapMap reference populations [30,31] with an r^2 threshold of 0.8; these were selected using the *tagger* application of the Genome Variation Server (SeattleSNPs Program for Genomic Applications PGA, 2009; http://gvs.gs.washington.edu/GVS/index.jsp). Because of the large size of *CARD11*, rather than choose tag SNPs, we genotyped SNPs approximately 8 kb apart. Second, we identified SNPs for fine mapping the chromosome 7 region. A 1-LOD drop region [32] of the chromosome 7 linkage peak reported by Stein et al. [16] was covered by genotyping SNPs approximately every 11 kb. For an SNP to be included on the custom array, it had to meet both a minor allele frequency threshold of 5% and SNP score quality criterion (Illumina SNP quality score > 0.6). SNP quality scores were determined by Illumina and are based on the probability of success of the assay and validation of the SNP in at least 2 populations.

Before analysis, Mendelian inconsistencies were removed using MARKERINFO (S.A.G.E. v 6.0). Call rates by plate were considered an additional measure of quality control and any SNPs that did not meet the call rate of 90% were excluded. Also, signal intensities were verified as falling into 3 distinct genotype groups, and thus no SNPs were lost because of inadequate signaling. Based on these quality control measures, a total of 119 SNPs were excluded from analyses (Supplementary Table 1). In addition, individual samples with call rates below 90% were excluded from the analysis.

2.3. Statistical analysis

Demographic factors included age, sex, HIV status, and the presence of Bacillus Calmette–Guérin (BCG) scar. Significant differences between individuals with and without TB were examined using the Pearson χ^2 for categorical variables and the Mann–Whitney test for age because age was nonnormally distributed. Before conducting association analysis, we examined the SNP genotypes for departure from Hardy–Weinberg proportions using Haploview [33].

Genetic association analyses were conducted using a generalized linear mixed model (GLMM) in SAS's PROC GLIMMIX procedure (SAS Institute, Cary, North Carolina), with each individual included in the analysis model. Because TB is a binary trait, we used a logit link function. To account for familial relationships, we used a generalized estimating equation framework and applied an exchangeable correlation structure within each pedigree, assuming that all individuals had a correlation of $\rho = 1$ with themselves and a correlation ρ with any other member in their family (*i.e.*, the G covariance matrix), defined as

$$G = \begin{pmatrix} 1 & \rho \\ & \ddots \\ \rho & 1 \end{pmatrix}. \quad (1)$$

Each SNP was entered into the GLMM individually, along with the 3 covariates, HIV status, sex, and dichotomized age (<5 years or ≥ 5 years), based on results by Lewinsohn et al. [34]. SNP genotypes were coded in an additive fashion to represent the number of minor (risk) alleles and were analyzed 1 at a time in the analysis, along with the aforementioned covariates. This generalized estimating equation model approach to examining genetic association to TB has been used in other family-based studies [35–37]. Also, this method allows for the inclusion of all individuals, both related and unrelated, into the analysis, thereby maximizing power [28],

In addition, we examined haplotypes in *IL6, CTSZ*, and *MC3R*; because of the large size and low SNP density in *CARD11*, we did not examine haplotypes for this gene. Based on an r^2 of 0.8, haplotypes were constructed in DECIPHER (S.A.G.E. v 6.0). Four haplotype blocks were identified in the analysis: 1 in *CTSZ*, 1 in *MC3R*, and 2 blocks in the *IL6* gene. The most likely haplotype phase was used for each individual in the analysis. Haplotypes were then treated similar to SNP genotypes in the GLMM analysis, coded in an additive fashion. Rare haplotypes with frequency $<10\%$ in the sample were pooled for analysis.

Finally, because some of our associated SNPs were identified in regions known to contain copy number variants (CNVs) [38], we investigated the Illumina intensity data for CNVs using PennCNV [39] and the default parameter settings. PennCNV infers DNA copy

Table 1
Descriptive statistics for analysis sample (% within category)

Category	Total (% of total sample)	Tuberculosis (TB)	No TB	*p* value[a]
Male	247 (44)	72 (29%)	175 (71%)	
Female	318 (56)	63 (20)	255 (80)	<0.0001
HIV negative	429 (83)	77 (18)	352 (82)	
HIV positive	91 (17)	58 (64)	33 (36)	0.0098
BCG scar	326 (67)	62 (19)	264 (81)	
No BCG scar[b]	140 (30)	45 (32)	95 (68)	0.0020
Total	565 (100)	135 (24)	430 (76)	
Age range		<1–53	<1–65	<0.0001[c]

[a]Pearson χ^2 test for difference in proportion of TB patients within category.
[b]BCG scar was unclear for some individuals, so numbers do not sum to 100%.
[c]Difference in age distributions in TB patients versus healthy individuals using the Mann–Whitney test.

number from SNP intensity data via a hidden Markov model, in which overall intensity and the individual allele intensities reflect an underlying, unknown copy number "state."

3. Results and discussion

After removal of samples with genotyping call rates less than 90%, 565 individuals with complete genotype data were included from both Phase I and Phase II. The sample comprised 318 females (56.3%) and 247 males (43.7%); 429 (82.5%) individuals were HIV negative, whereas 91 (17.5%) individuals were HIV positive, and the other individuals' HIV serostatuses were unknown (Table 1) because HIV testing was only conducted in adults and children who had HIV-positive mothers. A total of 135 (23.9%) individuals had culture-confirmed TB. Of the individuals without TB, 271 individuals either had a positive TST without TB disease at baseline or converted to TST positive sometime during study follow-up (48% of the total sample). The sample comprised 243 pedigrees, including 73 singletons, 230 parent–offspring pairs, and 32 sibling pairs, with a mean family size of 5.08 individuals and a standard deviation of 5.87. The median age was 16 years. TB patients were significantly older (median age = 28) than unaffected individuals (median age = 12; Mann–Whitney $p < 0.0001$).

Departure from Hardy–Weinberg equilibrium in the control individuals (without active TB) was found in 23 SNPs ($p < 0.0001$, data not shown). None of the candidate gene SNPs was significantly associated with TB susceptibility ($p > 0.10$). For the fine mapping analysis, a total of 1,479 SNPs were genotyped across the 17.84-Mb region on chromosome 7; of these, 1,367 met the 90% call rate threshold. No significant associations with TB susceptibility were reported with these SNPs ($p > 0.10$; Table 2). Both the candidate gene analysis and the fine mapping analysis were repeated using only the 429 HIV-negative individuals. Two SNPs, rs10233991 and rs12700594, were associated with TB susceptibility at the $\alpha = 0.05$ level, with nominal $p = 0.0308$ and $p = 0.0390$, respectively; rs12700594 is approximately 107 kb away from *CARD11*. Three additional SNPs had *p* values between 0.05 and 0.10. It is likely that more significant associations are seen in the HIV-negative subset because these individuals do not have a compromised immune system simply caused by HIV; in other words, HIV infection may confound the effects of genotype at certain loci. We could not examine the HIV-positive subset of individuals because of a reduced sample of only 91 such individuals, limiting our analysis.

Table 3
Results of *CTSZ* haplotype association analysis in entire sample and HIV-negative subsample

Haplotype (% of entire sample with haplotype)	Entire sample		HIV-negative subset	
	Odds ratio	*p* value	Odds ratio	*p* value
"GAGGG" (20.7)	1.5493	0.0281	1.3644	0.2436
All rare (pooled) (11.4)	1.6809	0.0354	2.1514	0.0128

CTSZ haplotype includes rs10369, rs9760, rs163790, rs163800, and rs163801.

We then conducted haplotype analyses in both the full sample and the HIV-negative subset. *CTSZ* haplotypes were constructed using the following SNPs: rs10369, rs9760, rs163790, rs163800, and rs163801. Haplotype analyses identified a significant association within the *CTSZ* gene (Table 3). The frequency of haplotype "GAGGG" was significantly greater in cases than in controls, adjusting for HIV status, age, and sex ($p = 0.0281$); this haplotype had a frequency of 21.0% in the full sample. Also, the number of copies of the rare haplotypes (frequencies less than 10%) were included in 1 indicator variable, and this variable was also significant ($p = 0.0354$). This suggests that at least 1 rare haplotype is associated with increased risk for TB, but because of our limited sample size, we did not have the statistical power to dissect which haplotype(s) was the risk haplotype(s); such examination of rare variants is a contemporary challenge in genetic epidemiology [40]. When only HIV-negative individuals were analyzed, the frequency of rare haplotypes was significantly greater in cases than in controls ($p = 0.0128$), although the GAGGG haplotype was no longer significantly associated ($p = 0.2436$). Haplotype analyses of *MC3R* and *IL6* did not yield statistically significant associations at the $\alpha = 0.05$ level (data not shown).

From these analyses, none of the candidate gene SNPs or fine mapping SNPs was significantly associated with TB susceptibility at a significance threshold accounting for multiple comparisons (for candidate genes, $p < 0.0125$, and for fine mapping, $p < 4.13 \times 10^{-5}$).

It is possible that significant SNP associations were not identified underlying the linkage peaks reported by Stein et al. [16] because the exact familial correlations could not be integrated into the generalized linear model; thus, correlations between full siblings were most likely underestimated. Also, linkage analysis is more powerful for detecting rare variants with strong effect sizes, whereas association analysis is more powerful for detecting common variants [41]; therefore, it is possible that our genome scan points to a rare variant (or multiple rare variants) on chromosome 7. It is also conceivable that the chromosome 7 linkage identified in our whole-genome scan was a false-positive finding, particularly because the trait status in 18 individuals changed [27].

Replication of a single-SNP between *CTSZ* and TB, as originally reported by Cooke et al. [21], was not successful. However, this

Table 2
Single-SNP association analysis results on chromosome 7 with $p < 0.10$ in entire sample and HIV-negative subsample

SNP	SNP location	Gene[a]	Entire sample		HIV-negative subset	
			Odds ratio	*p* value	Odds ratio	*p* value
rs10233991	170805		0.9983	0.36	0.9945	0.031
rs12700594	3123227		0.9988	0.51	0.9949	0.039
rs13437998	13838749		0.9991	0.61	0.9954	0.073
rs7811444	1562722	TMEM184A	0.9994	0.75	0.9960	0.083
rs7783310	1712658		0.9991	0.63	0.9957	0.093

[a]Genes containing SNPs within 500 kb were identified using SNPDoc (http://www.phs.wfubmc.edu/public/bios/gene/downloads.cfm).

significant association with *CTSZ* was based upon the analysis of a single SNP, rs34069356. This SNP did not meet design criteria for the GoldenGate assay and thus was not included in the panel of markers tested here. Thus, direct replication using published guidelines [42] of the Cooke et al. [21] analysis of this SNP was not possible, although we did identify nominally significant associations between *CTSZ* haplotypes and TB. It is possible that the risk allele reported by Cooke et al. [21] resides within these haplotypes. We could not evaluate this hypothesis via imputation, however, because rs34069356 is not included in the HapMap III set of SNPs. In general, our results support a role for *CTSZ* as a TB susceptibility gene, and further research in this area is warranted. Recently, a study in an independent Cape Town, South Africa study population both replicated the single SNP association of Cooke et al. and also found CTSZ haplotypes associated with TB (Adams et al. [49]). Future analyses will examine all of these *CTSZ* SNPs, including the SNP reported by Cooke et al. [21], in a larger, independent dataset.

The significant associations of the *MC3R* SNPs, rs3746619 and rs3827103, with TB susceptibility as reported by Cooke et al. [21] were also not replicated. Another potential reason for the nonreplication between the present study and that of Cooke et al. [21] could be differences in population genetics. The original association with *CTSZ* and *MC3R* was found in a South African population of fully independent sibling pairs, including "Coloreds," South Africans of mixed ancestry, but a case–control analysis in a West African sample within the same report failed to replicate the association. Based on our previous linkage disequilibrium (LD) and haplotype analyses, we found that the Ugandan population carries genetic distinctions from other African populations, including novel polymorphisms and LD structure in a different South African sample [43]. Therefore, it is not unexpected that the Cooke et al. [21] results were not replicated. In addition, we did not find associations with *IL6* as reported by Oral et al. [19]. IL-6 plays a role in TB immunopathogenesis, so the role of host genetics for this gene is not clear.

Our fine mapping results do not point to any clear candidates for causal variants. As seen in Table 2, many of these SNPs are not within 500 kb of a gene coding region. It is possible that these SNPs lie in regulatory regions, similar to those identified through expression quantitative trail loci (eQTL) studies [44]. Many trait-associated SNPs (>40%) are intergenic or intronic [45]; therefore, it is not surprising that we identified TB-associated SNPs in a potential gene desert. Interestingly, rs13437998 is approximately 1 megabase away from rs7787531, which was an imputed SNP that attained $p = 8 \times 10^{-6}$, reported by a recent genome-wide association analysis of TB in populations from Ghana and the Gambia [46]. Additionally, 2 SNPs, rs10233991 and rs13437998, reside in regions containing CNVs [38]; thus, we examined the Illumina intensity data for CNVs. We found that rs13437998 is in a CNV region with 3 copies (compared with the 2 expected when there are no CNVs; data not shown). Structural variants may be relevant for TB because other studies have demonstrated an association between *CCL3L1* copy number and HIV infection and progression [47]. CNVs may be in LD with SNPs; as a result, an observed SNP association may actually be caused by a CNV [48].

This study followed the previously reported linkage signals and their significant linkage to TB susceptibility in a genome-wide scan [16]. Therefore, levels of significance in these association tests should be interpreted with caution. The *p* values may not be the best determinant of significant association to TB susceptibility; with a purpose to locate and identify the most likely point in a region already identified by linkage, a nominal *p* value may be inappropriate given prior probability of association in that region. A major limitation to consider is the finite sample size of only 564 individuals in families, which was underpowered to detect small effect sizes. Another limitation is the spacing of the fine mapping SNPs; 11-kb spacing may not be dense enough to detect untyped loci of small effect in LD with our SNPs.

Acknowledgments

This work is supported in part by the National Institutes of Health, National Center for Research Resources Multidisciplinary Clinical Research Career Development Programs Grant (KL2RR024990), Tuberculosis Research Unit (Grants N01-AI95383 and HHSN266200700022C/N01-AI70022 from the National Institute of Allergy and Infectious Diseases), Grant R01HL096811 from the National Heart, Lung, and Blood Institute (NHLBI), and NHLBI Grant T32-HL0756. We acknowledge the invaluable contribution made by the study medical officers, health visitors, and laboratory and data personnel: Dr Lorna Nshuti, Dr Roy Mugerwa, Dr Christina Hirsch, Allan Chiunda, Brenda Okware, Mark Breda, Dennis Dobbs, Hussein Kisingo, Mary Rutaro, Albert Muganda, Richard Bamuhimbisa, Yusuf Mulumba, Deborah Nsamba, Barbara Kyeyune, Faith Kintu, Mary Nsereko, Gladys Mpalanyi, Janet Mukose, Grace Tumusiime, Pierre Peters, Alphonse Okwera, Keith Chervenak, Karen Morgan, Moses Joloba, Alfred Etwom, Micheal Angel Mugerwa, and Lisa Kucharski. We thank Dr Christopher Whalen for his instrumental role in the design of the household contact study. We also acknowledge and thank Dr Francis Adatu Engwau, Head of the Uganda National Tuberculosis and Leprosy Program, for his support of this project. We acknowledge the medical officers, nurses, and counselors at the National Tuberculosis Treatment Centre, Mulago Hospital, the Ugandan National Tuberculosis and Leprosy Program, and the Uganda Tuberculosis Investigation Bacteriological Unit, Wandegeya, for their contributions to this study. This study would not be possible without the generous participation of the Ugandan patients and families. Some of the results of this paper were obtained using the program package S.A.G.E., which is supported by a U.S. Public Health Service Resource grant (RR03655) from the National Center for Research Resources.

Appendix. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.humimm.2011.02.016.

References

[1] World Health Organization. Global Tuberculosis Control: Epidemiology, Strategy, Financing : WHO Report 2009. Geneva, World Health Organization, 2009.
[2] Murray CJ, Styblo K, Rouillon A. Tuberculosis in developing countries: burden, intervention and cost. Bull Int Union Tuberc Lung Dis 1990;65:6–24.
[3] Comstock G. Epidemiology of tuberculosis. Am Rev Respir Dis 1982;125:8–15.
[4] Skamene E, Schurr E, Gros P. Infection genomics: Nramp1 as a major determinant of natural resistance to intracellular infections. Annu Rev Med 1998;49:275–87.
[5] Kramnik I, Dietrick W, Demant P, Bloom B. Genetic control of resistance to experimental infection with virulent *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A 2000;97:8560–5.
[6] Kallmann F, Reisner D. Twin studies on the significance of genetic factors in tuberculosis. Am Rev Tuberc 1943;47:549–74.
[7] Simonds B. Tuberculosis in Twins. London, Pitman Medical 1963.
[8] Comstock G. Tuberculosis in twins: a re-analysis of the Prophit survey. Am Rev Respir Dis 1978;117:621–4.
[9] Shaw MA, Collins A, Peacock CS, Miller EN, Black GF, Sibthorpe D, et al. Evidence that genetic susceptibility to *Mycobacterium tuberculosis* in a Brazilian population is under oligogenic control: linkage study of the candidate genes NRAMP1 and TNFa. Tuber Lung Dis 1997;78:35–45.
[10] Moller M, de Wit E, Hoal EG. Past, present and future directions in human genetic susceptibility to tuberculosis. FEMS Immunol Med Microbiol 2010;58:3–26.
[11] Bellamy R, Beyers N, McAdam K, Ruwende C, Gie R, Samaai P, et al. Genetic susceptibility to tuberculosis in Africans: a genome-wide scan. Proc Natl Acad Sci USA 2000;97:8005–9.
[12] Greenwood C, Fujiwara T, Boothroyd L, Miller M, Frappier E, Fanning E, et al. Linkage of tuberculosis to chromosome 2q35 loci, including NRAMP1, in a large aboriginal Canadian family. Am J Hum Genet 2000;67:405–16.
[13] Jamieson S, Miller E, Black G, Peacock C, Cordell H, Howson J, et al. Evidence for a cluster of genes on chromosome 17q11–q21 controlling susceptibility to tuberculosis and leprosy in Brazilians. Genes Immun 2004;5:46–57.

[14] Miller E, Jamieson S, Joberty C, Fakiola M, Hudson D, Peacock C, et al. Genome-wide scans for leprosy and tuberculosis susceptibility genes in Brazilians. Genes Immun 2004;5:63–7.
[15] Stein CM, Zalwango S, Chiunda AB, Millard C, Leontiev DV, Horvath AL, et al. Linkage and association analysis of candidate genes for TB and TNFalpha cytokine expression: evidence for association with IFNGR1, IL-10, and TNF receptor 1 genes. Hum Genet 2007;121:663–73.
[16] Stein CM, Zalwango S, Malone LL, Won S, Mayanja-Kizza H, Mugerwa RD, et al. Genome scan of *M. tuberculosis* infection and disease in Ugandans. PLoS ONE 2008;3:e4094.
[17] van Crevel R, Ottenhoff T, van der Meek J. Innate immunity to *Mycobacterium tuberculosis*. Clin Microbiol Rev 2002;15:294–309.
[18] Ladel CH, Blum C, Dreher A, Reifenberg K, Kopf M, Kaufmann SH. Lethal tuberculosis in interleukin-6-deficient mutant mice. Infect Immun 1997;65: 4843–9.
[19] Oral HB, Budak F, Uzaslan EK, Basturk B, Bekar A, Akalin H, et al. Interleukin-10 (IL-10) gene polymorphism as a potential host susceptibility factor in tuberculosis. Cytokines 2006;35:143–7.
[20] Berrington WR, Hawn TR. *Mycobacterium tuberculosis*, macrophages, and the innate immune response: does common variation matter? Immunol Rev 2007; 219:167–86.
[21] Cooke GS, Campbell SJ, Bennett S, Lienhardt C, McAdam KP, Sow O, et al. Mapping of a novel susceptibility locus suggests a role for MC3R and CTSZ in human tuberculosis. Am J Respir Crit Care Med 2008;178:203–7.
[22] Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 2007;35:D61–5.
[23] Guwattude D, Nakakeeto M, Jones-Lopez E, Maganda A, Chiunda A, Mugerwa R, et al. Tuberculosis in household contacts of infectious cases in Kampala, Uganda. Am J Epidemiol 2003;158:887–98.
[24] Stein CM, Nshuti L, Chiunda AB, Boom WH, Elston RC, Mugerwa RD, et al. Evidence for a major gene influence on tumor necrosis factor-alpha expression in tuberculosis: path and segregation analysis. Hum Hered 2005;60:109–18.
[25] Diagnostic standards and classification of tuberculosis in adults and children. This official statement of the American Thoracic Society and the Centers for Disease Control and Prevention was adopted by the ATS Board of Directors, July 1999. This statement was endorsed by the Council of the Infectious Disease Society of America, September 1999. Am J Respir Crit Care Med;161:1376–95 2000.
[26] Targeted tuberculin testing and treatment of latent tuberculosis infection. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, July 1999: This is a joint statement of the American Thoracic Society (ATS) and the Centers for Disease Control and Prevention (CDC). This statement was endorsed by the Council of the Infectious Diseases Society of America (IDSA), September 1999, and the sections of this statement. Am J Respir Crit Care Med;161:S221–47 2000.
[27] Edwards BJ, Haynes C, Levenstien MA, Finch SJ, Gordon D. Power and sample size calculations in the presence of phenotype errors for case/control genetic association studies. BMC Genet 2005;6:18.
[28] Gray-McGuire C, Bochud M, Goodloe R, Elston RC. Genetic association tests: a method for the joint analysis of family and case-control data. Hum Genomics 2009;4:2–20.
[29] Stein CM. Genetic epidemiology of tuberculosis susceptibility: IMPACT of study design. PLoS Pathog 2011;7:e1001189.
[30] International HapMap Consortium. The International HapMap Project. Nature 2003;426:789–96.
[31] International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851–61.
[32] Mangin B, Goffinet B, Rebai A. Constructing confidence intervals for QTL location. Genetics 1994;138:1301–8.
[33] Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263–5.
[34] Lewinsohn DA, Zalwango S, Stein CM, Mayanja-Kizza H, Okwera A, Boom WH, et al. Whole blood interferon-gamma responses to *Mycobacterium tuberculosis* antigens in young household contacts of persons with tuberculosis in Uganda. PLoS ONE 2008;3:e3407.
[35] Velez DR, Wejse C, Stryjewski ME, Abbate E, Hulme WF, Myers JL, et al. Variants in Toll-like receptors 2 and 9 influence susceptibility to pulmonary tuberculosis in Caucasians, African-Americans, and West Africans. Hum Genet 2010; 127:65–73.
[36] Velez DR, Hulme WF, Myers JL, Stryjewski ME, Abbate E, Estevan R, et al. Association of SLC11A1 with tuberculosis and interactions with NOS2A and TLR2 in African-Americans and Caucasians. Int J Tuberc Lung Dis 2009;13: 1068–76.
[37] Velez DR, Hulme WF, Myers JL, Weinberg JB, Levesque MC, Stryjewski ME, et al. NOS2A, TLR4, and IFNGR1 interactions influence pulmonary tuberculosis susceptibility in African-Americans. Hum Genet 2009;126:643–53.
[38] Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, et al. Detection of large-scale variation in the human genome. Nat Genet 2004;36:949–51.
[39] Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF, et al. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res 2007; 17:1665–74.
[40] Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet 2008;40:695–701.
[41] Ardlie KG, Kruglyak L, Seielstad M. Patterns of linkage disequilibrium in the human genome. Nat Rev Genet 2002;3:299–309.
[42] Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, et al. Replicating genotype–phenotype associations. Nature 2007;447:655–60.
[43] Baker A, Randhawa A, Shey M, de Kock M, Kaplan G, Adams M, et al. Comparison of genotype frequencies in Toll-like receptor genes in Ugandans, South Africans, and African HapMap populations. Presented at the American Society of Human Genetics annual meeting, Honolulu, HI, 22 October 2009.
[44] Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet 2010;6;e1000888.
[45] Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A 2009;106:9362–7.
[46] Thye T, Vannberg FO, Wong SH, Owusu-Dabo E, Osei I, Gyapong J, et al. Genome-wide association analyses identifies a susceptibility locus for tuberculosis on chromosome 18q11.2. Nat Genet 2010;42:739–41.
[47] Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, et al. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. Science 2005;307:1434–40.
[48] McCarroll SA, Altshuler DM. Copy-number variation and association studies of human disease. Nat Genet 2007;39:S37–42.
[49] Adams LA, Moller M, Nebel A, Schreiber S, van der Merwe L, van Helden PD, Hoal EG. Polymorphisms in MC3R promoter and CTSZ 3'UTR are associated with tuberculosis susceptibility. Eur J Hum Genet 2011; doi:10.1038/ejhg2011.1.